A Time Series Study of Online Reviews for Experience Goods with Varying Text Lengths: A Case Study of Movie Reviews (Postprint)

Authors: Wang Jun, Li Zijian, Liu Xiaoman

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] This study classifies online reviews of experience goods into long-text and short-text categories based on text length, investigates the temporal and content characteristics of these two review types, and provides an intelligence basis for e-commerce platforms to understand patterns of consumer online review behavior and product demand preferences.

[Method/Process] Python web scraping technology is employed to collect relevant information on online reviews from movie review websites, construct time interval sequences of online reviews, and based on constructs from human behavior dynamics, examine the temporal characteristic patterns of posting behaviors for different review types. Text mining methods are utilized to identify the text content characteristics of different types of online reviews and conduct comparative analysis.

[Results/Conclusion] Using online reviews from movie review websites as the data source, this study summarizes from a temporal perspective that the time interval sequences of different types of online review behaviors follow a power-law distribution. From a text content perspective, it is found that the text content characteristics of different types of online reviews exhibit both certain similarities and significant differences.

Full Text

Preamble

Study on Time Series of Online Reviews for Experiential Products Based on Text Length: Taking Movie Reviews as an Example

Wang Jun, Li Zijian, Liu Xiaoman
School of Management, Jilin University, Changchun 130022

Abstract

[**Purpose/Significance**] This study categorizes online reviews of experiential products into long-text and short-text reviews based on text length, and explores the temporal and content characteristics of these two review types to provide an intelligence basis for e-commerce platforms to understand consumer online review behavior patterns and product demand preferences. [**Method/Process**] Using Python crawler language to obtain relevant information from online reviews on movie review websites, we construct online review time interval sequences. Based on concepts from human behavioral dynamics, we investigate the temporal characteristic patterns of different types of online review posting behaviors. Text mining methods are employed to identify content characteristics of different review types and conduct comparative analysis. [**Result/Conclusion**] Using online reviews from movie review websites as the data source, we find from a temporal perspective that the time interval sequences of different review types follow a power-law distribution. From a text content perspective, we discover that the textual content characteristics of different review types exhibit both similarities and significant differences.

Keywords: online reviews; time series analysis; text mining; content characteristics

1. Introduction

With the rapid development of e-commerce, the volume of online reviews has shown explosive growth, characterized by large spans in word count, numerous short-text reviews, and a power-law distribution between review length and quantity [1]. As the word count of online reviews increases, the semantic depth of the text strengthens, exerting a positive influence on review helpfulness [2]. From the perspective of content characteristics, short-text reviews feature simple content, single information expression, and short text length, requiring minimal time and effort from consumers to read. In contrast, long-text reviews contain relatively rich content with higher information density and longer text length, enabling consumers to obtain substantial product information and make decisions more easily [3]. In terms of review volume, short-text reviews far outnumber long-text reviews and are posted more intensively. Therefore, it is necessary to classify online reviews by text length and conduct comparative analysis between long-text and short-text reviews.

This study takes online reviews of experiential products as the research object, examining them from both time series and text content perspectives. By deeply mining long-text and short-text online reviews separately, we can help e-commerce platforms grasp consumer review patterns and assist consumers in obtaining needed information. Simultaneously, analyzing the behavioral patterns and content characteristics of online reviews with different text lengths holds important intelligence value for e-commerce platforms and regulatory authorities.

2. Literature Review

2.1 Temporal Dimension of Online Reviews

Existing online reviews generally feature timestamp functions, and this digital nature has gradually become an important dimension in online review research. Scholars have applied attribution theory using questionnaire surveys to conclude that shorter intervals between purchase and review posting can enhance consumers' perceived helpfulness of online reviews [4], and that the time interval between initial and follow-up reviews also affects this perception [5]. Zhang et al. used cluster analysis to divide the "initial-follow-up" review time intervals for search product reviews into stages and conducted text mining on review content within each stage [6]. Sun et al. found that longer intervals between movie trailer release and movie premiere lead to more positive consumer review sentiment [7]. L. Jin et al., from the perspectives of temporal distance theory and structural fit theory, argued that recent reviews have greater influence on consumers' near-term purchase decisions, while older reviews become relatively more influential for long-term decisions [8].

2.2 Text Mining of Online Reviews

Text mining technology is widely applied in online review research, primarily involving feature extraction and sentiment analysis. Typical machine learning algorithms for text feature extraction include Support Vector Machine (SVM) models [9], Hidden Markov Model (HMM) [10], and Conditional Random Fields (CRFs) models [11], with some scholars also using complex network theory for text content analysis [12]. Dong et al. used text mining technology to extract content features from online reviews across three e-commerce platforms and analyzed their correlations [13]. Zhang et al. extracted information features from mobile phone user reviews, finding that higher product update frequency correlates with more positive consumer evaluations [14].

Product feature extraction is closely related to sentiment analysis, with most current research using text mining for sentiment analysis of online reviews [15]. Li et al. used convolutional neural networks for sentence-level sentiment classification of online reviews [16]. Ma et al. used the ROST EA text analysis tool to conduct sentiment analysis of review content, established scoring, and built regression models to examine correlations between scoring and content evaluation [17]. Zheng et al. combined semantic and statistical methods for sentiment classification to construct a sentiment ontology for analyzing sentiment polarity and intensity in online reviews [18]. Wei et al. used multivariate regression analysis on sales data from 100 automotive brands (2012-2016) and found that online review sentiment polarity directly affects brands and indirectly influences product sales [19]. K. Y. Lee et al. collected real reviews from Amazon and through text mining discovered that online review sentiment polarity negatively correlates with consumer product acceptance [20].

2.3 Online Review Helpfulness

Current research on online review helpfulness focuses on using mathematical methods to construct evaluation models for ranking review usefulness. For example, Guo et al. built a review helpfulness model for O2O online reviews based on fuzzy analytic hierarchy process [21]. Xiu et al. evaluated online review usefulness from an information transmission efficiency perspective based on communication patterns and barriers theory [22]. Other studies examine helpfulness influencing factors, such as exploring review usefulness from experiential versus search product perspectives [23]. Fang et al. constructed a econometric model of online review helpfulness influencing factors, confirming that brand reputation and product type moderate review helpfulness [24]. Wang et al. used experimental methods to prove that different temporal and social distances affect users' evaluation of online review helpfulness [25]. Some scholars have introduced cognitive psychology concepts, such as Wang et al., who used eye-tracking technology to visualize consumer attention to online reviews and identify factors affecting perceived review helpfulness [26].

Online reviews exhibit enormous volume, low average word count, and extremely unbalanced text length distribution. In most studies, numerous short reviews are filtered out due to low word count and information density. However, short-text reviews, despite their massive quantity, still contain important intelligence value in aggregate. Existing research mostly classifies online reviews by time, paying insufficient attention to differences in volume, sentiment polarity, and temporal distribution across different text lengths. Few studies conduct comparative analysis of online reviews categorized by text length from both temporal characteristics and text mining dimensions. Therefore, this study uses text length as the classification criterion, dividing online reviews into long-text and short-text categories, and employs text mining methods combined with the digital indicator of review timestamps to conduct in-depth analysis from both time series and content perspectives, aiming to reveal consumer online review behavior patterns and content characteristics, thereby enriching research on experiential product online review features.

3. Research Design

3.1 Research Framework

The research framework is shown in Figure 1 [Figure 1: see original paper].

3.1.1 Crawling and Classifying Online Review Information

Based on Python programming language, we develop a web crawler to capture relevant information from target product online reviews, including review content, posting time, user rating, and user level. Using text statistical tools, we count the word count of review texts and specifically classify online reviews by character count into long-text and short-text reviews. Following the classification rules of "Douban Movies," another major Chinese movie review website, which categorizes reviews above 140 characters as professional reviews and those at or below 140 characters as non-professional short reviews, this study adopts the same rule: reviews exceeding 140 characters are classified as information-rich long-text reviews, while those at or below 140 characters are classified as information-sparse short-text reviews. Based on the captured timestamp information, these two types of reviews are arranged in chronological order of posting, yielding long-text review time series {A1, A2, A3…An} and short-text review time series {B1, B2, B3…Bn}. We then calculate the time intervals between adjacent reviews to obtain new time interval sequences {C1, C2, C3…Cn-1} and {D1, D2, D3…Dn-1}.

3.1.2 Characterizing Temporal Features of Consumer Review Behavior

We describe the dynamic changes in structured information such as review quantity, sentiment tendency, and ratings across different time dimensions for different text types. Based on theories and methods of temporal measurement in human behavioral dynamics, we introduce relevant indicators including time intervals, power-law exponents, and burstiness to systematically characterize the temporal interval features of review posting, further analyzing the temporal characteristic patterns of consumer review behavior.

3.1.3 Text Content Mining Based on Review Length

We examine the distribution patterns of high-frequency words in different review types and compare the similarities and differences in review content from three perspectives: character features, emotional expression features, and movie content features.

3.2 Research Object Selection

This study selects movie reviews as the research object, primarily because existing research focuses on experiential products (books, movies) and search products (cameras, mobile phones) [24]. Experiential products, due to their experiential and intangible nature, make it difficult for consumers to obtain relevant information before consumption [27], leading to higher risks and uncertainties in purchase decisions. To avoid information asymmetry risks, consumers need to read online reviews of such products [28]. Compared to search product reviews, consumers are more dependent on experiential product reviews, highlighting the importance of studying experiential product online reviews. Movies, as typical experiential products, align well with this research context.

Regarding platform selection, this study uses “Maoyan Movie,” a third-party platform, as the data source. Originally “Meituan Movie,” launched by Meituan in February 2012 and renamed in January 2013, Maoyan has become a widely used movie application with high market share, generating massive volumes of consumer movie reviews that provide important data support for this study.

3.3 Data Collection and Preprocessing

We use a self-developed Python crawler program to capture 117,342 online review data entries for the popular movie *Hidden Man* from Maoyan Movie. The captured information includes review content, posting time, user level, and user rating. The data collection period spans 153 days from the movie's premiere on July 13, 2018, to December 13, 2018.

4. Case Analysis

4.1 Dynamic Characteristics of Online Reviews

From a time series perspective, we analyze the dynamic changes in long-text and short-text reviews from two dimensions: review quantity and sentiment tendency.

4.1.1 Dynamic Characteristics of Review Quantity Analyzing both review types from the perspective of review quantity, we use weekly, daily, and hourly time units to count review data from the first 12 weeks after the movie's release, categorized by review type.

Weekly Analysis
Figures 2 [Figure 2: see original paper] and 3 [Figure 3: see original paper] describe the changing trends of different review types over time. Overall, both long-text and short-text review quantities show consistent decreasing trends: the first week has the most reviews and the most active commenting behavior, the second week sees a sharp decrease, followed by stabilization at low levels. For long-text reviews, the first week contains 1,829 reviews, accounting for 81.36% of the total, while the second week has 326 reviews, representing an 82.12% decrease from the first week. For short-text reviews, the first week contains 62,633 reviews (54.47% of the total), while the second week has 21,542 reviews, a 65.61% decrease. Overall, review quantity shows a decreasing trend because early movie reviews occur during high-heat periods with substantial word-of-mouth dissemination. As time passes, decreasing consumer attention leads to weakened word-of-mouth behavior, reducing movie review heat and failing to stimulate potential consumers, resulting in fewer posted reviews.

Daily Analysis
Using daily time units, we examine the distribution patterns of long-text and short-text reviews across the seven days of a week. For comparative analysis, we standardize the review quantities in each statistical interval, with results shown in Figure 4 [Figure 4: see original paper]. The trends of both review types are consistent: quantities remain low from Monday to Thursday, increase significantly from Friday, peak on Saturday, with Friday-Sunday being the weekly peak period. The difference is that long-text reviews are more concentrated during peak periods and show more dramatic weekly fluctuations than short-text reviews.

Hourly Analysis
Using hourly time units, we divide the 24-hour day into four periods: 0:00-6:00 (early morning), 6:00-12:00 (morning), 12:00-18:00 (afternoon), and 18:00-24:00 (evening). After standardizing the distribution of different review types across 24 hours, the results are shown in Figure 5 [Figure 5: see original paper]. Both review types show consistent daily trends, with peaks at 12:00, 17:00, and 22:00, reaching the daily maximum at 22:00. However, differences exist: between 22:00 and 10:00 the next day, long-text reviews are posted more intensively, while short-text reviews are more concentrated during other periods. This is because daytime hours are typically fragmented, leading consumers to post short-text reviews, while evenings provide continuous time for editing long-text reviews, making consumers more inclined to post long-text reviews at night.

4.1.2 Dynamic Characteristics of Sentiment Tendency Ratings represent consumer satisfaction levels—generally, higher ratings reflect higher satisfaction and more positive sentiment. Therefore, this study uses ratings to measure sentiment tendency. We calculate the average ratings and standard deviations of long-text and short-text reviews for weeks 1-5 after the movie's premiere to analyze review extremity, as shown in Tables 1 and 2 . Both review types show common characteristics and certain differences in rating variance distribution. Horizontally, both variances decrease as time since premiere increases, because longer screening times reduce attention and discussion intensity. Vertically, short-text reviews maintain consistently higher rating variance across all time intervals than long-text reviews, indicating more intense discussion, higher movie heat, greater controversy, and stronger stimulation of consumer curiosity. Long-text reviews provide more comprehensive movie evaluations with lower emotional extremity and discussion intensity.

From a rating perspective, the gradual overall rating increase represents reputation recovery. From the perspective of review extremity corresponding to ratings in each time interval, movie review ratings and extremity show an inverse relationship: within the same time interval, higher ratings correspond to lower review extremity. Consumers are more enthusiastic about giving high scores and more positive evaluations when movie review extremity is low, while controversial movies receive lower scores and more negative sentiment.

4.2 Temporal Interval Characteristics of Online Reviews

4.2.1 Distribution Characteristics of Review Time Intervals Time interval refers to the time difference between adjacent reviews. If the total online review time sequence contains N data points, the time interval sequence contains N-1 data points. To investigate the temporal characteristic patterns of long-text and short-text online review behaviors, we introduce the concept of time intervals from online human behavioral dynamics and use power-law fitting for characterization. Table 3 describes basic features of online review time interval sequences, including minimum, maximum, mean, and standard deviation values, using "minutes" as the basic unit.

The shortest time interval in both long-text and short-text review sequences is 0 minutes. However, the proportion of reviews with 0-minute intervals in short-text reviews far exceeds that in long-text reviews, with significant differences in maximum, mean, and standard deviation values. To further compare these review types, we plot scatter diagrams using logarithmic transformations of time interval sequences and review quantities. Some observation points represent data with low frequency and long intervals, showing obvious "fat-tail" phenomena at the distribution ends. Using least squares method to fit the main data after removing these outliers (Figures 6 [Figure 6: see original paper] and 7 [Figure 7: see original paper]), both review types achieve goodness-of-fit above 90%, confirming that both long-text and short-text online review time intervals follow power-law distributions.

4.2.2 Index Characteristics of Review Time Intervals To further explore consumer posting behavior characteristics for different text lengths, we introduce the burstiness coefficient from online human behavioral dynamics, combined with review quantity and power-law exponents, to describe behavioral patterns.

Burstiness is a statistic describing the intensity of human activity in short periods and silence in long periods, calculated as [29]:

$$B = \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau}$$

where σ_τ is the standard deviation of time interval sequence τ, and m_τ is its mean. For exponential distributions where standard deviation equals mean, B equals 0. More pronounced "fat-tail" distributions indicate larger differences between σ_τ and m_τ, making B approach 1. Using formula (1), we calculate burstiness coefficients of 0.5797 for long-text reviews and 0.7598 for short-text reviews.

As Table 4 shows, long-text and short-text review quantities differ significantly: short-text reviews number 114,994 (98% of total), while long-text reviews number only 2,348 (2%). This is because long-text reviews require more time and effort to compose, resulting in fewer postings. Both review types show obvious burstiness—during the movie's initial release, topicality and heat generate large volumes of reviews in short periods. Short-text review time intervals follow a power-law distribution with exponent -1.803, while long-text reviews follow -1.218. The substantial difference in exponents arises because long-text reviews are posted more sporadically with generally longer intervals between adjacent reviews, while short-text reviews are posted more intensively with higher proportions of short intervals and lower proportions of long intervals, resulting in a smaller absolute power-law exponent for long-text reviews.

5. Text Content Analysis

To mine and summarize content characteristics of long-text and short-text reviews, we use Python's jieba segmentation package to extract and analyze keywords from review content through the following process:

1. **Construct stopword list**: Remove meaningless characters such as letters, emoticons, timestamps, and function words like "的" (de), "地" (de), "然后" (then), "后来" (later) during segmentation.
2. **Sort by word frequency**: Extract keywords after stopword removal, count occurrence frequencies, and sort in descending order.
3. **Categorize high-frequency words by features**: Based on distribution patterns and frequency attributes of high-frequency words across review types, we reflect content differences. We categorize keywords like "actor," "director," "dialogue," "acting," and "actor names" as character features; words describing consumer emotions ("like," "good," "average," "worthwhile," "trash," "boring") as emotional expression features; and "plot," "style," "storyline," "ending" as movie content features.

High-frequency words in online movie reviews reflect reviewers' primary concerns—more frequent high-frequency words of a certain feature attribute indicate greater audience attention to that feature. Based on all review information from 153 days after the movie's premiere, we compile the top 120 high-frequency words for each review type, visualize them as word clouds where font size corresponds to frequency (Figures 8 [Figure 8: see original paper] and 9 [Figure 9: see original paper]), and present the top 50 high-frequency words in Table 5 .

Long-text and short-text reviews show both obvious commonalities and significant differences, reflecting that consumer groups corresponding to the two review types have partially different product concerns. From the distribution proportions of character, emotional, and movie content features, long-text reviews prioritize movie content features (50%), followed by character features (28%), with emotional expression features accounting for only 22%. Short-text reviews allocate 40% to emotional expression features, 34% to movie content features, and 26% to character features.

Across both review types, 12 identical high-frequency words appear in character features (24% of total), 9 in emotional expression features (18%), and 10 in movie content features (20%), yielding an overall high-frequency word similarity of 62%. Regardless of review type, attention to character features remains consistent at 28% and 26% respectively. Among all high-frequency words, "actor names" appear most frequently, indicating significant star effects in movie reviews. The most frequent word in character features is "Jiang Wen"; in emotional expression features, "like"; and in movie content features, "story" and "plot."

The key difference is that short-text reviews show the highest proportion of emotional expression features (40%), indicating that consumers posting such reviews focus more on emotional expression. Emotional expression features include both positive and negative words ("like," "good," "average," "worthwhile," "bad," "trash"), consistent with earlier analysis that short-text reviews exhibit higher emotional extremity and less consensus. Long-text reviews focus more on the movie itself—movie content features—and primarily use positive and neutral emotional words ("like," "humorous," "good," "why"), resulting in generally higher ratings and lower emotional extremity.

6. Discussion of Research Findings

Using Python crawler language to capture online review information, we classify reviews into long-text and short-text categories based on character count. By introducing relevant indicators from human behavioral dynamics theory, we analyze the distribution patterns of both review types from the posting time perspective. Results show that both long-text and short-text review time interval sequences follow power-law distributions with typical "fat-tail" features, with exponents of -1.218 and -1.803 respectively, and both exhibit strong burstiness. Based on large-sample online review data, we analyze dynamic patterns of long-text and short-text reviews from perspectives of review quantity, sentiment polarity, and ratings, conduct text content mining, compare content characteristics, and reveal behavioral patterns and textual features of online reviews with different text lengths.

The practical implication is that e-commerce platforms can segment consumer markets based on online review text length, which helps understand consumer posting behavior characteristics and product content preferences, enabling targeted marketing strategies. For consumers inclined to post long-text reviews, platforms should emphasize product content features; for short-text review posters, they should use emotionally charged words and topics to attract attention and discussion. By understanding review quantity and temporal characteristics, platforms can timely grasp changes in review types and content, and intervene accordingly—for example, encouraging consumers to post long-text reviews more promptly to increase their quantity and shorten posting intervals when long-text reviews are insufficient. From the platform optimization perspective, besides sorting reviews by recency, reviews could be categorized by "product features," "emotional expression features," and "content features" using text mining technology to help consumers quickly find needed information, reduce decision-making time costs, and alleviate information overload.

This study has several limitations. First, it only uses movie reviews as samples of experiential product reviews without incorporating other experiential product reviews, limiting sample representativeness. Second, using rating variance to measure emotional extremity may reduce analytical accuracy. Third, using only one third-party platform as the data source limits generalizability. Future research should verify whether online review posting intervals on other e-commerce platforms follow power-law distributions, exhibit strong burstiness, and fall within similar power-law exponent ranges.

References

[1] Zhang L, Qian G, Fan W, et al. Research on sentiment analysis of lightweight reviews [J]. Journal of Software, 2014, 25(12): 2790-2807.

[2] Mudambi SM, Schuff D. What makes a helpful online review? A study of customer reviews on Amazon.com [M]. Mount Laurel: Society for Information Management and The Management Information Systems Research Center, 2010.

[3] Chevalier JA, Mayzlin D. The effect of word of mouth on sales: online book reviews [J]. Journal of marketing research, 2006, 43(3): 345-354.

[4] Wang T, Wang K, Chen H. When can time intervals enhance the perceived usefulness of online reviews: Based on attribution theory [J]. Journal of Business Economics, 2015, 280(2): 46-56.

[5] Hu C, Ning C. When are online follow-up reviews more useful than initial reviews?—Based on the moderating effect of time interval and product type [J]. Forecasting, 2017, 36(4): 36-42.

[6] Zhang Y, Peng L, Hong C. Empirical study on time series correlation characteristics of online user follow-up review behavior—Taking mobile phone review data from JD.com as an example [EB/OL]. [2019-01-11]. http://kns.cnki.net/kcms/detail/11.1762.G3.20181030.0921.004.html.

[7] Sun C, Liu Y. The impact of online movie trailer release on box office—Based on text sentiment analysis method [J]. Chinese Journal of Management Science, 2017, 25(10): 151-160.

[8] Jin L, Hu B, He Y. The recent versus the Out-Dated: an experimental examination of the time-variant effects of online consumer reviews [J]. Journal of retailing, 2014, 90(4): 552-566.

[9] Liu J, Zhang P, Lu Y. Automatic identification of messages related to adverse drug reactions from online user reviews using feature-based classification [J]. Iranian journal of public health, 2014, 43(11): 1519-1527.

[10] Borrajo L, Seara Vieira A, Iglesias EL. TCBR-HMM: an HMM-based text classifier with a CBR system [J]. Applied soft computing, 2015, 26(1): 463-473.

[11] Chasin R, Woodward D, Witmer J, et al. Extracting and displaying temporal and geospatial entities from articles on historical events [J]. The computer journal, 2014, 57(3): 403-426.

[12] Sun Z, Gu J, Yang J. Chinese entity relation extraction method based on deep learning [J]. Computer Engineering, 2018, 44(9): 164-170.

[13] Dong S, Wang X, Ge Z. Analysis of online review content features on B2C shopping websites based on text mining [J]. Library Theory and Practice,

2017(6): 54-58.

[14] Zhang L, Wu F, Huang L. Product innovation research based on user network review information [J]. Soft Science, 2015, 29(5): 12-16.

[15] Liu M, Wang X, Li H, et al. Sentiment analysis of network product reviews based on text mining [J]. Journal of Liaoning University of Technology (Natural Science Edition), 2018, 38(5): 330-335.

[16] Li J, Li H. Short text review product feature extraction and sentiment classification research based on deep learning [J]. Information Studies: Theory & Application, 2018, 41(2): 143-148.

[17] Ma S, Xu X. Research on user online evaluation based on review sentiment analysis—Taking Douban movies as an example [J]. Library and Information Service, 2016, 60(10): 95-102.

[18] Zheng L, Wang H. Sentiment polarity and intensity analysis of online reviews based on sentiment ontology [J]. Journal of Management Sciences, 2015, 28(4): 85-93.

[19] Wei R, Zheng J. Research on emotional marketing effects of online reviews [J]. Journal of Shanghai University of International Business and Economics, 2018, 25(4): 72-80.

[20] Lee KY, Yang SB. The role of online product reviews on information adoption of new product development professionals [J]. Internet research, 2015, 25(3): 435-452.

[21] Guo S, Zhang X, Li Z. Research on mobile O2O online review helpfulness ranking model oriented to user information needs—Taking Meituan as an example [J]. Library and Information Service, 2015, 59(23): 85-93.

[22] Xiu G, Wang J, Guo S. Online review usefulness evaluation introducing information transmission efficiency [J]. Information Science, 2019, 37(1): 43-50.

[23] Jiang X. What kind of product reviews are most helpful?—Research on the influence of online review quantitative and textual features on helpfulness [J]. Foreign Economics & Management, 2015, 37(4): 41-55.

[24] Fang J, Wang Y, Zhao Z. The moderating effect of different product brand reputation on influencing factors of online review helpfulness [J]. Soft Science, 2016, 30(3): 108-112.

[25] Wang J, Ding D. Research on the relationship between online review helpfulness and temporal and social distance [J]. Information Studies: Theory & Application, 2016, 39(2): 73-77, 81.

[26] Wang C, Gao H. Research on influencing factors of online review helpfulness perception with follow-ups—Based on eye-tracking experiments [J]. Journal of Modern Information, 2018, 38(12): 70-77, 90.

[27] Cheema A, Papatla P. Relative importance of online versus offline information for Internet purchases: product category and Internet experience effects [J]. Journal of business research, 2010, 63(9): 979-985.

[28] Dellarocas C, Zhang X, Awad NF. Exploring the value of online product reviews for forecasting sales: The case of motion pictures [J]. Journal of interactive marketing, 2007, 21(4): 23-45.

[29] Goh KI, Barabasi AL. Burstiness and memory in complex systems [J]. EPL (Europhysics Letters), 2008, 81(4): 48002-p1-48002-p5.

Author Contributions

Wang Jun: Supervised research direction;
Li Zijian: Data analysis and main paper writing;
Liu Xiaoman: Data collection and processing.

Abstract: [Purpose/significance] According to text length, online experiential product reviews are divided into long-text online reviews and short-text online reviews. Exploring the temporal and content characteristics of these two types provides an intelligence basis for e-commerce platforms to understand consumer online review behavior patterns and product demand preferences. [Method/process] Python crawler language is employed to collect information from online reviews on movie review websites, construct online review time interval sequences, and based on human behavioral dynamics theory, investigate temporal characteristic patterns of different review types. Text mining methods are used to discover and compare content characteristics across review types. [Result/conclusion] Using movie review website data, the study concludes that time interval sequences obey power-law distributions across different review types. From a text mining perspective, content characteristics show both similarities and significant differences.

Note: Figure translations are in progress. See original paper for figures.